



Assessment of Stereoscopic Display Systems for Assisting in Route Clearance Manipulation Planning Tasks

by Andrew S. Bodenhamer

ARL-TR-4195

August 2007

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

DESTRUCTION NOTICE—Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Aberdeen Proving Ground, MD 21005-5425

ARL-TR-4195**August 2007**

Assessment of Stereoscopic Display Systems for Assisting in Route Clearance Manipulation Planning Tasks

Andrew S. Bodenhamer
Human Research & Engineering Directorate, ARL

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) August 2007		2. REPORT TYPE Final		3. DATES COVERED (From - To) August 2006	
4. TITLE AND SUBTITLE Assessment of Stereoscopic Display Systems for Assisting in Route Clearance Manipulation Planning Tasks				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Andrew S. Bodenhamer (ARL)				5d. PROJECT NUMBER 62716AH70	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory Human Research & Engineering Directorate Aberdeen Proving Ground, MD 21005-5425				8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-4195	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBERS	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This study was conducted to objectively compare how the use of a three-dimensional (3-D) or two-dimensional (2-D) visual display affects manipulation planning performance in a spatial perception task relevant to the operation of the Buffalo vehicle manipulator arm. Portions of the results are also generalizable to any tele-operated precision manipulation task. The task evaluated involves judging the position of the Buffalo arm relative to targets and obstacles as seen in the visual display from the arm camera. Thirty-two Soldiers trained to use the route clearance vehicle participated in this study at Fort Leonard Wood, Missouri, during the summer of 2006. Significant planning performance benefits were found when the 3-D visual display was used as opposed to the 2-D display. A significant correlation between subjective confidence ratings and objective performance was found with 3-D displays but not with 2-D displays. Most participants indicated that they preferred using the 3-D display instead of the 2-D display.					
15. SUBJECT TERMS stereoscopic vision; robotics; teleoperation; route clearance					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES 30	19a. NAME OF RESPONSIBLE PERSON Andrew S. Bodenhamer
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED			19b. TELEPHONE NUMBER (Include area code) (573) 329-8705

Contents

List of Figures	iv
List of Tables	iv
1. Introduction	1
2. Method	3
2.1 Participants	3
2.2 Apparatus.....	3
2.3 Experimental Design	6
2.4 Procedure.....	7
3. Results	8
4. Discussion	10
5. Conclusions	12
6. References	13
Appendix A. Experiment Scenarios	15
Appendix B. Participant Comments	21
Distribution List	22

List of Figures

Figure 1. Buffalo mine-protected vehicle.	1
Figure 2. Stereoscopic cameras.....	4
Figure 3. Experiment work station.....	5
Figure 4. Mean score by view mode (error bars show 95% confidence interval).	8
Figure 5. Mean confidence ratings by view mode (error bars show 95% confidence interval). ..	9
Figure 6. Response ratio by confidence rating and view mode.	10
Figure A-1. Terminal screen capture for scenario 1.	15
Figure A-2. Terminal screen capture for scenario 2.	15
Figure A-3. Terminal screen capture for scenario 3.	16
Figure A-4. Terminal screen capture for scenario 4.	16
Figure A-5. Terminal screen capture for scenario 5.	17
Figure A-6. Terminal screen capture for scenario 6.	17
Figure A-7. Terminal screen capture for scenario 7.	18
Figure A-8. Terminal screen capture for scenario 8.	18
Figure A-9. Terminal screen capture for scenario 9.	19
Figure A-10. Terminal screen capture for scenario 10.	19
Figure A-11. Terminal screen capture for scenario 11.	20
Figure A-12. Terminal screen capture for scenario 12.	20

List of Tables

Table 1. Questions for video set A.....	6
Table 2. Questions for video set B.....	7
Table B-1. Participant comments.....	21

1. Introduction

This study is one of a series of stereovision studies conducted through the Robotics Collaboration Army Technology Objective (RC ATO). The objectives of this research initiative were to identify human visual perception shortfalls in current and future Army robotic systems, assess the maturity of commercially available equipment to address these issues, and integrate and test stereovision systems to quantify performance improvements. The scope of this investigation has, to date, primarily focused on tele-operated ground robots and manipulators used by the U.S. Army engineer, infantry, and explosive ordnance disposal Soldiers.

The Buffalo is a heavy mine protected vehicle that is currently being used in operations in theater for route clearance missions. It has been widely used in Operation Iraqi Freedom for improvised explosive device (IED) searches and as a command and control platform for mine-clearing operations. The Buffalo is a blast-resistant vehicle intended to protect Soldiers from the effects of mine blasts. It has a tele-operated hydraulic arm used for identifying suspected mines and IEDs. The operator for the tele-operated arm is situated inside the vehicle in the front right passenger seat. In order for the operator to successfully identify and classify suspected mines and IEDs, s/he must have the ability to see the object with great clarity. A camera is mounted on the arm with the images displayed to the operator inside the vehicle. Figure 1 shows the Buffalo vehicle with arm partially extended.



Figure 1. Buffalo mine-protected vehicle.

The operator's only tool to aid in the identification of an unknown object is the "fork" attached to the end of the boom arm. With that tool, the operator must very carefully peel away something covering the object to get a better view of it, nudge some debris to the side to clear some of the area around the object, or rake some of the surface dirt, possibly uncovering some tell-tale signs of an IED. All these manipulations must be done with precision in case an IED is encountered. This task is primarily visual, and the operator relies heavily on the camera mounted near the end of the manipulator arm. It was this factor that raised the issue of whether three-dimensional (3-D) vision might be an aid in interrogating (investigating) an unknown object. Two fundamental questions needed to be answered to meet the objectives of the RC ATO as delineated earlier in this section. First, would 3-D produce more information or a different quality of information for the operator, which could make the operator's task of deciding if an object is truly an IED an easier task or a more accurate task? Second, how could we quantify the result of using stereopsis to show that its effect is worthy of consideration for the operator's task? Simply demonstrating user preference is not usually sufficient for initiating costly changes in a system.

Prior studies by Cole, Merrit, Fore, and Lester (1990) and Drascic (1991), which regard stereoscopic vision capabilities, support the concept that benefits could be observed for an operator performing tasks similar to those needed to perform the route clearance mission with the Buffalo manipulator arm. However, the manipulation tasks evaluated in these studies are not fully generalizable to those required of the route clearance mission. A preliminary study was conducted by U.S. Army Research Laboratory researchers in December 2005 to observe expert Buffalo operators during IED investigation and to allow them to subjectively evaluate differences between two-dimensional (2-D) and 3-D displays for performing the IED investigation task. Pettijohn, Vaughan, and Bodenhamer (2005) indicate that 3-D may provide benefit for the precise manipulation of the Buffalo arm and improve confidence in identifying IEDs. Yet, objective performance data were still not available.

From observation of tele-operated manipulation, it can often be seen that even a moderately skilled operator generally employs a closed loop feedback method when s/he is performing a unique manipulation task. This follows a pattern such as 1) small manipulator movement, 2) visual assessment, 3) new movement vector decision. This is likely because of one or more factors inherent to tele-operated manipulation (e.g., control activation to end effector movement lag time, imprecise manipulator movement, or imperfect visual perception). Of particular interest is the cognitive aspect of this task, the visual assessment and subsequent movement decision, collectively referred to in this study as "manipulation planning".

The objective of this particular research effort is to objectively compare how the use of a 3-D or 2-D visual display affects manipulation planning performance in a spatial perception task that is relevant to the operation of the Buffalo arm and generalizable to any tele-operated precision manipulation. The task involves judging the position of the Buffalo arm relative to targets and obstacles as seen in the visual display from the arm camera. The task of visually evaluating if and how a collision (intentional or unintentional) will occur between the arm and an object is

especially important in the use of the Buffalo to provide maximal dexterity of a very powerful and heavy manipulator without damaging the arm or accidentally detonating an IED. The hypothesis is that 3-D view mode can improve operator manipulation planning performance over 2-D view mode. This is tested through answering the following questions. Is planning performance in 3-D significantly different than 2-D? Are operators significantly more confident when using 3-D than when using 2-D? How does confidence relate to planning performance in both 2-D and 3-D?

2. Method

2.1 Participants

Thirty-two Soldiers were recruited to participate in this study. These Soldiers were from a mixture of active Army, Army reserve, and National Guard units that were students in the route clearance vehicle operator's course at Fort Leonard Wood, Missouri. The study was conducted during the final two days of the 14-day course, so that all the participants were fully trained and had some degree of experience with operating the Buffalo in field exercises.

Participants' ages ranged from 18 to 46 years, with a median age of 23. All the participants were male. Participant rank ranged from Private (E-1) to Sergeant First Class (E-7), with a median rank of Specialist (E-4). Twenty-seven participants were military occupational specialty (MOS) 21B (Combat Engineer), three were MOS 21E (Heavy Construction Equipment Operator), and one was MOS 21N (Construction Equipment Supervisor). All participants were verified to have normal visual acuity (20/30 or better), stereo depth perception, and color vision in both eyes. The voluntary, fully informed consent of the persons used in this research was obtained as required by 32 Code of Federal Regulations 219 (OSD, 1999) and Army Regulation (AR) 70-25 (HQDA, 1990). The investigators have adhered to the policies for the protection of human subjects as prescribed in AR 70-25.

2.2 Apparatus

This study did not involve interactive use of the Buffalo vehicle or manipulator arm but was conducted with pre-recorded video from a December 2005 study of initial integration of stereoscopic camera systems onto the Buffalo mine protected vehicle. This procedure allowed the tasks to be uniform across all participants and avoided inherent complications attributable to the varying skill levels for the participants directly controlling the arm.

Video used in the study was initially recorded from a pair of Panasonic color video cameras mounted on the arm of the Buffalo, as seen in figure 2. The distance between the centers of the left and right camera lenses was approximately equal to a mean human inter-pupillary distance of 64 mm. Each camera had a field of view of approximately 39 by 30 degrees. The stereo camera

pair was converged at a range of approximately 1 meter and statically fixed in relation to the end effector (the Buffalo fork).



Figure 2. Stereoscopic cameras.

The left and right video camera signals (L- and R-channels) were combined by a video multiplexer that creates alternating video fields of L- and R-channel images and records to a digital video tape. Upon playback, this composite interlaced signal is transmitted to the video display as a sequence of alternating, L- and R-channel, full-screen images.

The video contains manipulations performed by expert operators, a noncommissioned officer and civilian who are U.S. Army Engineer School trainers for the Buffalo. The scenarios depicted in the video were selected and set up by subject matter experts from the U.S. Army Counter-Explosive Hazards Center. Following the stereoscopic experimental design guidance of Merritt (1988), efforts were made to ensure that a fair comparison was made between the 2-D and 3-D display modes and that tasks evaluated were representative of the operational environment. Although the materials used in the scenarios may differ from those encountered in the current theater of operations, the manipulations needed to interrogate (investigate) the target sites were consistent with the current route clearance mission. The environmental conditions when the video was recorded were such that there was diffuse sunlight and no sustained wind. The video has been edited to 12 short clips that show the scene, with the arm moving, for a period between 15 and 30 seconds. The clip ends when there is an imminent occlusion or contact of the forks of the Buffalo arm and an object in the scene. Obvious clues as to whether the forks will collide or miss the object are not contained in the preceding video segment. At the end of each clip, the video pauses for 30 seconds to allow the participant to answer questions associated with that clip. For the 3-D clips, the video paused in 3-D. There was also a 15-second delay (blue screen) between each clip to allow the participant to

be informed of what was to be shown in the next clip. The video clips were originally recorded in 3-D (interlaced), but duplicate 2-D copies of each were made with the use of StereoMovie Maker¹ version 0.93 for Windows². Both the 2-D and 3-D versions of the video were presented at 720 x 480 resolution at 30 frames per second. It is important to recognize that despite having the same display resolution, the perceived information contained in the 2-D and 3-D videos is different. Since the stereo images were transmitted and stored via field sequential multiplexing, the two images are correlated (as opposed to being totally independent channels of information), and the perceived loss of resolution is 0.707. When StereoMovie Maker generated a 2-D image by reconstituting a full video frame from one field by duplicating lines without new information, the perceived loss of resolution for the 2-D videos is 0.500. Screen captures (2-D) of the terminal view of the 12 clips are shown in appendix A. This is the view, after the brief motion, when the participant was asked the question associated with each video clip.

For 2-D and 3-D trials, the video was played on a Pavonine Dimen³ G170S 17-inch 2-D/3-D liquid crystal display monitor connected to a personal computer. The operator wore linearly polarized glasses to see the 3-D video display. The polarized glasses allow only the L- or R-channel at a time to be seen by only the left or right eye. The interlaced left and right camera images of the properly spaced video cameras are seen as one combined L- and R-eyed image that the brain translates into a stereoscopic image of the scene. Polarized glasses were not worn during the 2-D trials. The experiment work station was a table and chair set up in an enclosed, climate-controlled building, as seen in figure 3.

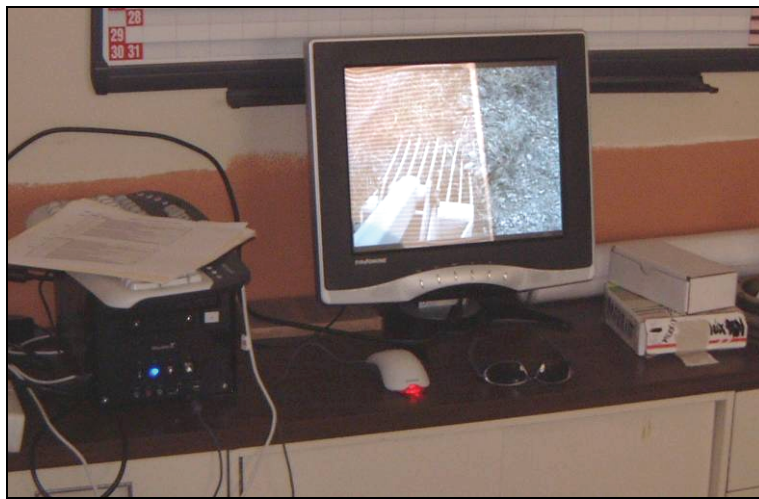


Figure 3. Experiment work station.

¹StereoMovie Maker is a registered trademark of Masuji SUTO (not an acronym).

²Windows is a registered trademark of Microsoft Corporation.

³Dimen is a registered trademark of Pavonine, Inc.

2.3 Experimental Design

This experiment was a single factor within-subjects design. The independent variable was display type, 2-D or 3-D. The dependent variables were perception responses and confidence rating. These dependent variables and their measures are described in more detail following.

During the 30-second pause at the end of each video clip, the participant was asked to answer a question associated with that clip. The responses were from a given set of two or three responses. The final measure was the score for each group of trials (number of correct responses/total number of evaluations).

Participants were also asked to provide a rating of their confidence in their answer. This confidence was expressed on a scale from 1 to 5 where 1 represents “very unsure/guess” and 5 represents “very confident.”

The 12 clips were sorted into two groups of six; each group had the same composition of motion types and similar objects. These motion types were “swing right,” “swing left,” and “extend”. Copies (2-D and 3-D) of each clip were sorted to make two video sets so that no participants will see the same event in both 2-D and 3-D groups. Within each group in each set, the video order was randomized. One set begins with 2-D and the other begins with 3-D. Balancing the order of the video clips within each video set was not performed because of the limitations of displaying the video on a tape device and the aggregate score structure of the data collection. Tables 1 and 2 display an ordered list of the questions asked and associated answers for each clip in each set.

Table 1. Questions for video set A.

02	2D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the box? Answer: hit
01	2D – If you continue to extend the arm will the tips of the forks hit the <u>top</u> , <u>side</u> or <u>overshoot</u> the box? Answer: top
03	2D – If you continue to swing the arm to the right will the tips of the forks go <u>over</u> or <u>under</u> the bench top? Answer: over
06	2D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the logs? Answer: hit
04	2D – If you swing the arm right will you <u>hit</u> or <u>miss</u> the microwave? Answer: miss
05	2D – If you continue to extend the arm will the tips of the forks <u>hit</u> the log or go <u>underneath</u> ? Answer: underneath
09	3D – If you continue to swing the arm to the right will the tips of the forks hit the microwave <u>body</u> , <u>door</u> , or <u>miss</u> altogether? Answer: door
10	3D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the rock? Answer: hit
11	3D – If you swing the arm to the right will the tips of the forks go <u>over</u> or <u>under</u> the bench? Answer: under
07	3D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the microwave? Answer: hit
12	3D – If you continue to extend the arm will the tips of the forks <u>hit</u> or go <u>under</u> the bag? Answer: hit
08	3D – If you continue to extend the arm will the tips of the forks hit the <u>top log</u> , the <u>bottom log</u> , or go <u>beneath</u> the log pile? Answer: beneath

Table 2. Questions for video set B.

02	3D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the box? Answer: hit
03	3D – If you continue to swing the arm to the right will the tips of the forks go <u>over</u> or <u>under</u> the bench top? Answer: over
01	3D – If you continue to extend the arm will the tips of the forks hit the <u>top</u> , <u>side</u> or <u>overshoot</u> the box? Answer: top
04	3D – If you swing the arm right will you <u>hit</u> or <u>miss</u> the microwave? Answer: miss
06	3D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the logs? Answer: hit
05	3D – If you continue to extend the arm will the tips of the forks <u>hit</u> the log or go <u>underneath</u> ? Answer: underneath
11	2D – If you swing the arm to the right will the tips of the forks go <u>over</u> or <u>under</u> the bench? Answer: under
07	2D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the microwave? Answer: hit
10	2D – If you swing the arm left will you <u>hit</u> or <u>miss</u> the rock? Answer: hit
12	2D – If you continue to extend the arm will the tips of the forks <u>hit</u> or go <u>under</u> the bag? Answer: hit
09	2D – If you continue to swing the arm to the right will the tips of the forks hit the microwave <u>body</u> , <u>door</u> , or <u>miss</u> altogether? Answer: door
08	2D – If you continue to extend the arm will the tips of the forks hit the <u>top log</u> , the <u>bottom log</u> , or go <u>beneath</u> the log pile? Answer: beneath

2.4 Procedure

Volunteers received an overview of the experiment, details of the procedures, and information about any risks involved with their participation. The volunteers read and signed an informed consent form if they wished to participate. The participants then completed a brief demographics questionnaire. Participants completed visual acuity, stereo depth, and color deficiency tests using a Titmus⁴ vision screening device.

Each participant was given a familiarization session to become comfortable with the task of evaluating the video and with viewing the 3-D display. The participant was shown 2-D and 3-D version of 2 “training” clips. These video clips lasted for approximately 3 minutes and consisted of a variety of manipulations with the Buffalo arm using the same camera view as was to be presented during the experiment. The participants were told that at the end of each clip, the video would pause and they would be asked about a decision for moving the arm; they would then be expected to base their answer on their perception of the motion or positioning of the arm during the clip. They were informed that the question set would include items such as

“If I swing the arm to the right/left, will the fork hit or miss the _____?”

“If I keep extending the arm on its current trajectory, will the fork hit or miss the _____?”

Participants also received instruction how to give a confidence rating for each of their responses. Participants were not allowed to watch others perform the experiment.

⁴Titmus is a registered trademark of Titmus Optical, Inc.

3. Results

SPSS⁵ for Windows, Release 13, was used for statistical analysis. As illustrated in figure 4, the mean score for 2-D view mode was 0.474 (standard deviation [SD] = 0.180), and for 3-D view mode was 0.672 (SD = 0.167). The skewness (2-D: -0.324, standard error [SE] = 0.414; 3-D: -0.273, SE = 0.414) and kurtosis (2-D: 0.823, SE = 0.809; 3-D: 0.289, SE = 0.809) were within ± 2 times their standard errors so that it can be assumed that the data are normally distributed and parametric tests are appropriate. Criteria for statistical tests of significant differences were set at $\alpha = 0.05$.

If it was assumed that participants randomly chose their answer from the two or three options presented, the expected score based upon these random guesses is 0.445. A one-sample t-test was performed to check for significant difference from the expected mean. For 2-D view mode $T(31) = 0.909$, $p = .370$ and for 3-D view mode $T(31) = 7.705$, $p < .001$. Thus, the mean score in 3-D view mode was significantly different from the expected mean, while 2-D was not. In other words, participants did not score significantly better in 2-D view mode than if they had made random guesses for each trial.

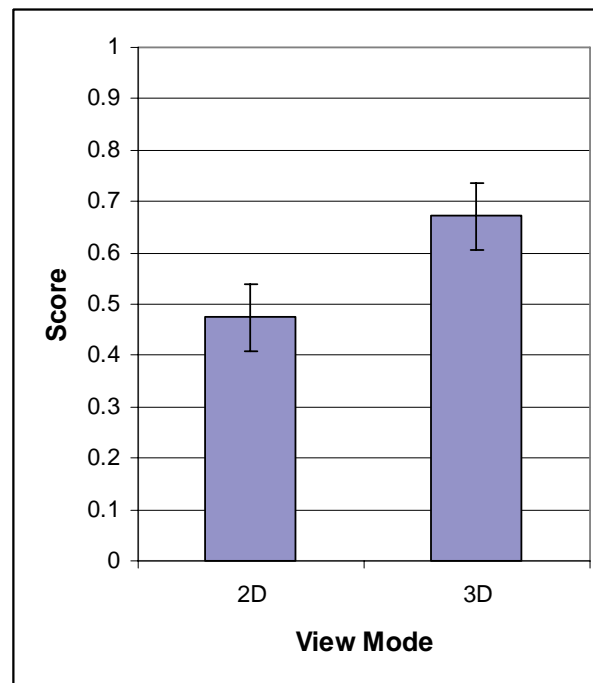


Figure 4. Mean score by view mode (error bars show 95% confidence interval).

⁵SPSS, which stands for Statistical Package for the Social Sciences, is a registered trademark of SPSS, Inc.

Next, a paired sample, within-subjects t-test was performed to identify whether an individual can be expected to have significantly different scores for 2-D and 3-D view modes. It was found that the difference in scores is significant, $T(31) = -4.381, p < .001$. Individuals assessed the scene and made manipulator motion decisions better in 3-D than in 2-D view mode.

A Pearson correlation of scores within subjects found $r = -.085, p = .644$. An individual's performance in 2-D does not correspond with the magnitude of the increase of score when s/he is performing in 3-D view mode across a range of target scenarios.

Repeating these tests but paired by target rather than participant, we find that a significant difference exists as well, $T(11) = -2.912, p = .014$. Targets similar to those presented in this study can be assessed better in 3-D than in 2-D view mode. Furthermore, a Pearson correlation of scores within targets yields $r = .757, p = .004$. There is a significant positive correlation between how well a target can be assessed to make a manipulation decision in 3-D compared to 2-D by a group of individual operators.

Confidence ratings (figure 5) for answers made during 2-D view mode had a mean of 3.91 (SD = 0.473) and during 3-D view mode had a mean of 4.23 (SD = 0.419).

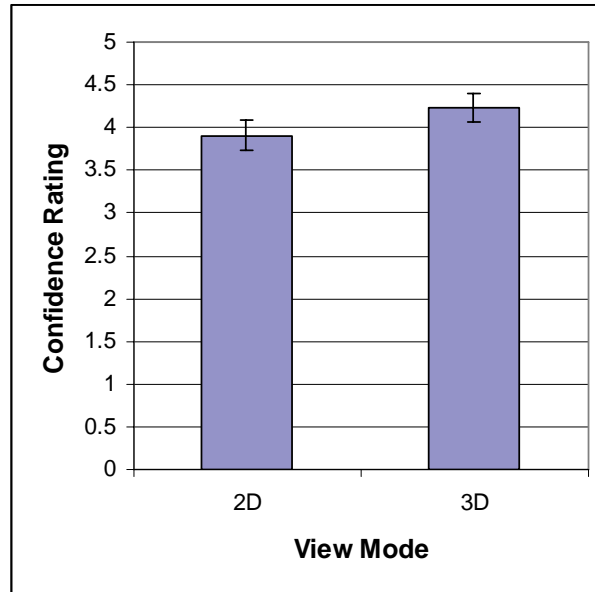


Figure 5. Mean confidence ratings by view mode (error bars show 95% confidence interval).

A non-parametric test was used to compare the confidence ratings for the two view modes. The paired samples, within-subjects Wilcoxon Signed Ranks Test found that $W(31) = -2.870, p = .004$. Thus, confidence ratings were significantly higher for 3-D than 2-D view mode.

To further investigate the relation of confidence ratings and view mode, given a perceived confidence rating, what is the likelihood of the movement decision being correct or incorrect? In other words, how well do confidence ratings predict manipulator planning task performance? This

is plotted in figure 6 as the ratio of correct to incorrect responses presented by confidence rating and view mode.

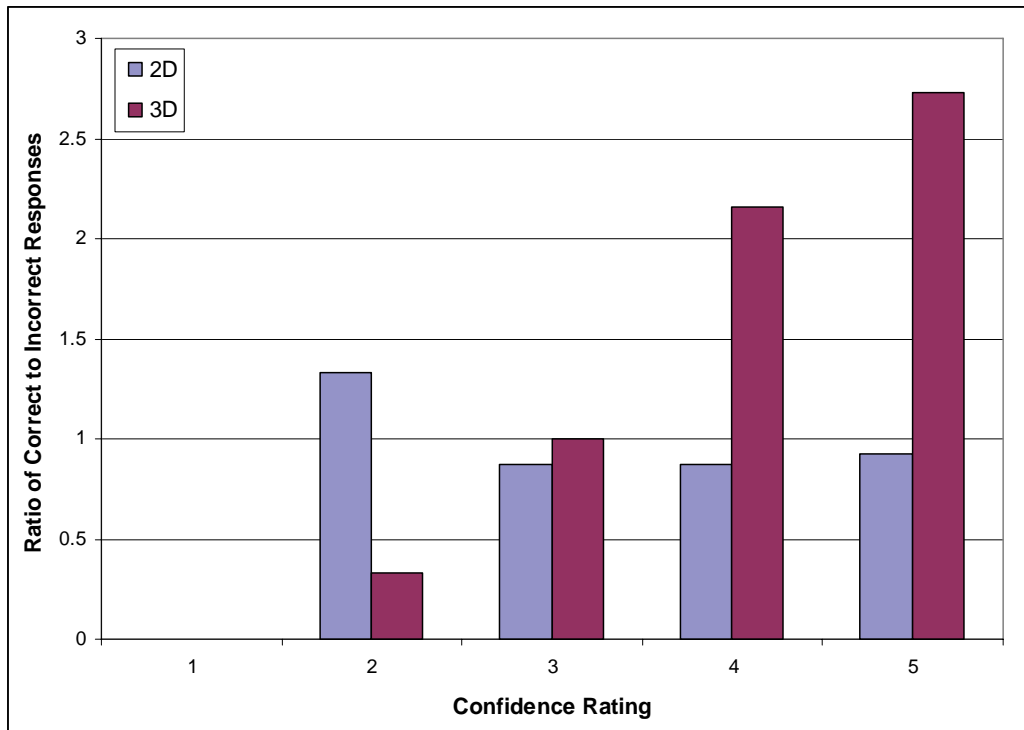


Figure 6. Response ratio by confidence rating and view mode.

Spearman's non-parametric correlation of confidence rating and response ratio shows that in 2-D view mode, $\rho = .400$, $p = .505$ and in 3-D view mode, $\rho = 1.00$, $p < .001$. The perception of confidence in 2-D view mode could not be shown to have meaning, while in 3-D view mode, the perception of confidence significantly corresponds to the odds of successful decision making.

After completing all trials, participants were asked, "How do you feel about the difference between 2-D and 3-D video when performing this type of task?" A full list of their responses is presented in appendix B. 22 Soldiers expressed an overall favorable opinion of using 3-D as opposed to 2-D. Five Soldiers preferred 2-D or had an unfavorable opinion regarding 3-D. Five Soldiers did not submit a response.

4. Discussion

The hypothesis that 3-D view mode can improve operator manipulation planning performance is supported by the results. Scores were significantly higher in 3-D than 2-D, as were confidence ratings. The magnitude of this improvement was an additional 19.8% correct manipulation decisions. The practical question that remains to be answered is how does this translate into an

improvement in reducing interrogation time and reduction in materiel lost to accidental activation of explosive devices or other damage?

The fact that 2-D decision making was not significantly different from random guesses and that confidence in 2-D was meaningless is of particular concern. What factors allow current operators to successfully perform the mission in the real world? Efforts were taken to make this comparison fair: using trained operators with recent experience and giving as much as 30 seconds of moving video for each scenario to allow motion parallax cues to be present. One explanation may be that most operators use shadows as a primary feedback in lieu of depth perception. The video in this study was nearly shadow free because of diffuse lighting conditions. This could bias the results in favor of 3-D, but direct overhead lighting should not be an expected condition for the performance of tele-operated manipulation in a tactical environment. On-board lighting for a robotic system may not provide the needed shadow cues either.

Another explanation is that current robotic manipulators are too slow to necessitate instantaneous manipulator planning. Most current manipulator controllers use a “joint control” scheme and only allow the activation of 1 or 2 degrees of freedom at any given time. This is an inherently slow process, allowing time for the closed loop perception cycle to account for the decrease in the ability to make instantaneous manipulation planning decisions.

One particularly interesting finding that was not an explicit research question was that an individual’s performance in 2-D did not correlate with the magnitude of the increase in score when s/he is performing in 3-D view mode across a range of target scenarios. For example, those who scored poorly in 2-D mode were not also the lowest scorers in 3-D mode (and vice versa). Individuals did not tend to simply experience a set percentage of performance “boost” when switching from manipulator planning in 2-D to 3-D. There are individual differences that make someone “good” at perception in 2-D or 3-D but not always both.

The most important finding within the context of this study is the data showing that 3-D vision does not simply enhance the manipulation decision-making ability but provides an entirely new capability (the ability of the operator to use his or her inherent feelings of confidence to improve task performance). In 2-D, however, an operator who is fully confident in his or her perception is often wrong. Not only does this seem inefficient, but if the operator is not aware of this deficiency, then during certain circumstances, this handicap could lead to accidents and is likely a source of fatigue and frustration to the operator. Furthermore, despite the potential benefits, there remains a question of whether the awareness of this confidence correlation when one is operating in 3-D mode could lead to unacceptable risk-taking behavior. An ensuing experiment may be warranted to fully investigate how and why the performance-confidence correlation fails in both 2-D and 3-D modes, since there may be confounding factors not addressed in this study.

5. Conclusions

The results of this experiment indicate that stereo-vision systems do have a performance benefit for an operator when s/he is encountering a unique tele-operated manipulation task relevant to the current Army tactical environment. Although the current availability and maturity of field-ready stereo-vision systems leave much to be desired, there is a strong continuous need within the Army for better robots and remote manipulators.

As manipulators become more advanced (such as six-degree-of-freedom end effector control or semi-autonomous manipulation) and users demand more efficient task performance, system designers will have to address the lack of proprioceptive feedback and stereoscopic vision that people take for granted in their own advanced manipulator (human arm) usage. Future research should investigate how autonomous assistance and stereoscopic vision systems could provide a key portion of the user interface to develop an intuitive, versatile, and efficient means to perform remote manipulation.

6. References

- Cole, R. E.; Merritt, J. O.; Fore, S.; Lester, P. Remote manipulator tasks impossible without stereo TV. *SPIE* Vol. 1256, Stereoscopic Displays and Applications, 1990.
- Drasic, D. Skill acquisition and task performance in teleoperation using monoscopic and stereoscopic video remote viewing. *Proceedings of the Human Factors Society 35th Annual Meeting*, 1991.
- Headquarters, Department of the Army. *Use of Volunteers as Subjects of Research*; AR 70-25; Washington, DC, 1990.
- Merritt, J. O. Often-overlooked advantages of 3-D displays. *SPIE* Vol. 902, Three-Dimensional Imaging and Remote Sensing Imaging, 1988.
- Office of the Secretary of Defense. Protection of Human Subjects, 32 Code of Federal Regulations, Part 219, Washington, DC, 1999.
- Pettijohn, B.A.; Vaughan, B.D.; Bodehamer, A.S. Personal communication, U.S. Army Research Laboratory: Fort Leonard Wood, MO, December 22, 2005.

INTENTIONALLY LEFT BLANK

Appendix A. Experiment Scenarios

Note: Color, contrast, and resolution of the following screen captures are not identical to the video as seen by the participants.



Figure A-1. Terminal screen capture for scenario 1.



Figure A-2. Terminal screen capture for scenario 2.



Figure A-3. Terminal screen capture for scenario 3.



Figure A-4. Terminal screen capture for scenario 4.



Figure A-5. Terminal screen capture for scenario 5.



Figure A-6. Terminal screen capture for scenario 6.



Figure A-7. Terminal screen capture for scenario 7.



Figure A-8. Terminal screen capture for scenario 8.



Figure A-9. Terminal screen capture for scenario 9.



Figure A-10. Terminal screen capture for scenario 10.



Figure A-11. Terminal screen capture for scenario 11.



Figure A-12. Terminal screen capture for scenario 12.

Appendix B. Participant Comments

The following is a list of comments when participants were asked, “How do you feel about the difference between 2-D and 3-D video when performing this type of task?”

Twenty-two Soldiers expressed an overall favorable opinion of using 3-D as opposed to 2-D. Five Soldiers preferred 2-D or had an unfavorable opinion regarding 3-D. Five Soldiers did not submit a response.

Table B-1. Participant comments.

Big Difference. Prefer 3-D
3-D means less time on target and more confident
3-D gives me a headache. But 3-D helps with extending
Not much better in 3-D.
I liked 3-D more. More confident in 3-D.
3-D is more complicated. I can tell depth slightly better in 2-D.
3-D was better. Could be useful in theater.
I like 3-D better.
I like 3-D.
3-D is better.
3-D is better.
3-D was difficult, it played with my eyes. It felt less concrete and bothered my eyes.
3-D is nice but I would like to switch back and forth.
I was more confident with 3-D. Would have to touch ground using 2-D.
3-D made me a little nauseous.
3-D was better. Much more confident with 3-D.
3-D was a little annoying. 2-D was tough compared to 3-D.
3-D was better. This could make a big difference in theater.
3-D was better.
More confident in 3-D.
Large difference with 3-D.
Liked 3-D, but it takes time to get used to.
3-D is better. Takes some time to get used to it.
3-D was easier.
3-D is better.
3-D has a good benefit.
I liked 2-D, hard to adjust to 3-D.

NO. OF
COPIES ORGANIZATION

1 DEFENSE TECHNICAL
(PDF INFORMATION CTR
ONLY) DTIC OCA
8725 JOHN J KINGMAN RD
STE 0944
FORT BELVOIR VA 22060-6218

1 US ARMY RSRCH DEV & ENGRG CMD
SYSTEMS OF SYSTEMS
INTEGRATION
AMSRD SS T
6000 6TH ST STE 100
FORT BELVOIR VA 22060-5608

1 DIRECTOR
US ARMY RESEARCH LAB
IMNE ALC IMS
2800 POWDER MILL RD
ADELPHI MD 20783-1197

1 DIRECTOR
US ARMY RESEARCH LAB
AMSRD ARL CI OK TL
2800 POWDER MILL RD
ADELPHI MD 20783-1197

2 DIRECTOR
US ARMY RESEARCH LAB
AMSRD ARL CS OK T
2800 POWDER MILL RD
ADELPHI MD 20783-1197

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR M DR M STRUB
6359 WALKER LANE SUITE 100
ALEXANDRIA VA 22310

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR ML J MARTIN
MYER CENTER RM 2D311
FT MONMOUTH NJ 07703-5601

5 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MZ A BODENHAMER
199 E 4TH ST STE C TECH PARK BLDG 2
FT LEONARD WOOD MO 65473-1949

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MD T COOK
BLDG 5400 RM C242
REDSTONE ARSENAL AL 35898-7290

NO. OF
COPIES ORGANIZATION

1 COMMANDANT USAADASCH
ATTN AMSRD ARL HR ME A MARES
5800 CARTER RD
FT BLISS TX 79916-3802

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MO J MINNINGER
BLDG 5400 RM C242
REDSTONE ARSENAL AL 35898-7290

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MM DR V RICE-BERG
BLDG 4011 RM 217
1750 GREELEY RD
FT SAM HOUSTON TX 78234-5094

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MG R SPINE
BUILDING 333
PICATINNY ARSENAL NJ 07806-5000

1 ARL HRED ARMC FLD ELMT
ATTN AMSRD ARL HR MH C BURNS
BLDG 1467B ROOM 336
THIRD AVENUE
FT KNOX KY 40121

1 ARMY RSCH LABORATORY - HRED
AVNC FIELD ELEMENT
ATTN AMSRD ARL HR MJ D DURBIN
BLDG 4506 (DCD) RM 107
FT RUCKER AL 36362-5000

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MK MR J REINHART
10125 KINGMAN RD
FT BELVOIR VA 22060-5828

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MV HQ USAOTC
S MIDDLEBROOKS
91012 STATION AVE ROOM 348
FT HOOD TX 76544-5073

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MY M BARNES
2520 HEALY AVE STE 1172 BLDG 51005
FT HUACHUCA AZ 85613-7069

1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MP D UNGVARSKY
BATTLE CMD BATTLE LAB
415 SHERMAN AVE UNIT 3
FT LEAVENWORTH KS 66027-2326

NO. OF
COPIES ORGANIZATION

- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MJK J HANSBERGER
JFCOM JOINT EXPERIMENTATION J9
JOINT FUTURES LAB
115 LAKEVIEW PARKWAY SUITE B
SUFFOLK VA 23435
- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MQ M R FLETCHER
US ARMY SBCCOM NATICK SOLDIER CTR
AMSRD NSC SS E BLDG 3 RM 341
NATICK MA 01760-5020
- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MY DR J CHEN
12423 RESEARCH PARKWAY
ORLANDO FL 32826
- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MS MR C MANASCO
SIGNAL TOWERS 118 MORAN HALL
FORT GORDON GA 30905-5233
- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MU M SINGAPORE
6501 E 11 MILE RD MAIL STOP 284
BLDG 200A 2ND FL RM 2104
WARREN MI 48397-5000
- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MF MR C HERNANDEZ
BLDG 3040 RM 220
FORT SILL OK 73503-5600
- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MW E REDDEN
BLDG 4 ROOM 332
FT BENNING GA 31905-5400
- 1 ARMY RSCH LABORATORY - HRED
ATTN AMSRD ARL HR MN R SPENCER
DCSFDI HF
HQ USASOC BLDG E2929
FORT BRAGG NC 28310-5000
- 1 ARMY G1
ATTN DAPE MR B KNAPP
300 ARMY PENTAGON ROOM 2C489
WASHINGTON DC 20310-0300

NO. OF
COPIES ORGANIZATION

- 1 DR THOMAS M COOK
ARL-HRED LIAISON
PHYSICAL SCIENCES LAB
PO BOX 30002
LAS CRUCES NM 88003-8002
- 1 US ARMY AVIATION AND MISSILE
RD&E CTR WEAPONS SCIENCES DIR
ATTN AMSRD AMR WS PL JIM KIRSCH
REDSTONE ARSENAL AL 35898-5000
- 1 THE MERRITT CORPORATION
ATTN JOHN MERRITT
82 SOUTH ST POB 728
WILLIAMSBURG MA 01096
- ABERDEEN PROVING GROUND
- 1 DIRECTOR
US ARMY RSCH LABORATORY
ATTN AMSRD ARL CI OK TECH LIB
BLDG 4600
- 1 DIRECTOR
US ARMY RSCH LABORATORY
ATTN AMSRD ARL CI OK TP S FOPPIANO
BLDG 459
- 1 DIRECTOR
US ARMY RSCH LABORATORY
ATTN AMSRD ARL HR MR F PARAGALLO
BLDG 459